



Proyecto entorno a el problema de estimación de provisiones en la industria aseguradora empleando el método Chain-ladder

Santiago Prieto Betancur

Contents

1	Fase de comprensión del negocio/problema	3
1.1	Objetivos del negocio/problema	3
1.2	Evaluación de la situación	4
1.3	Objetivos en la ciencia de datos	5
1.4	Plan de proyecto	5

1 FASE DE COMPRENSIÓN DEL NEGOCIO/PROBLEMA

En el marco de la metodología de desarrollo de proyectos de CRISP-DM (Cross-Industry Standard Process for Data Mining) es necesario establecer en un primer momento el entendimiento del negocio o problema a tratar el cual empleará herramientas del análisis de datos.

De esta manera, en el problema de estimación de provisiones (reservas por pérdidas) en la industria aseguradora el objetivo principal es destinar un fondo de reservas o de aprovisionamientos de una cartera con el fin de liquidar aquellos productos de seguros reportados por los clientes después de un determinado siniestro. Por lo tanto, en el problema de estimación de provisiones el intento es estimar las reclamaciones de seguros que debe atender una empresa aseguradora pero que aun no han sido reportados y proyectar adicionalmente las ultimas cantidades de perdida en las cuales podría incurrir una compañía atendiendo estos siniestros. Adicionalmente, es de gran importancia entender que este problema al tratarse de una cuestión de naturaleza predictiva implica que una de las suposiciones subyacentes en la solución de esta incógnita es que los desarrollos de patrones que explican las perdidas históricas en las estimación de estas provisiones son un indicador para desarrollar patrones para perdidas futuras.

1.1 OBJETIVOS DEL NEGOCIO/PROBLEMA

A medida que el mercado asegurador se expande y más compañías incursionan en este negocio es importante desarrollar e implementar mecanismos que hagan rentable la oferta de seguros accediendo a un amplio número de clientes y a la par manteniendo una oferta de productos competitivos en el mercado. Es por esto, que se hace indispensable desarrollar un plan de trabajo en el cual se cumplan los siguientes objetivos:

- Recopilar, organizar y analizar datos relacionados con las pólizas a pagar después del reporte de un siniestro por los clientes de una compañía aseguradora.
- Determinar patrones o modelos que permitan reservar los montos adecuados de aprovisionamiento para los gastos operativos en una compañía aseguradora.
- Optimizar las reservas de cartera apresadas con el fin de mantener un flujo de capital mayor en inversiones o financiación de la compañía aseguradora y a su vez estimar el capital necesario para cubrir la liquidación de las pólizas reportadas.

De esta manera, para la consecución de este proyecto es necesario de la cooperación de distintos departamentos dentro de la estructura organizacional de una compañía de seguros

como lo son por lo general el comité de manejo del riesgo, el comité de inversión, el comité de auditoría y los asesores en el área de actuaría y IT (Information Technology) ya que de esta manera es posible articular mejor la información y a su vez tener acceso a datos de interés en el estudio y permitir igualmente su correcto procesamiento y análisis.

1.2 EVALUACIÓN DE LA SITUACIÓN

Para este proyecto es necesario disponer de una base de datos que permita determinar si el modelo predictivo a desarrollar brinda buenos resultados al momento de fijar las reservas de provisiones en una compañía aseguradora. Es por esto, que en el trabajo a desarrollar se empleará la base de datos CAS Loss Reserve la cual fue construida a partir de la base de datos Schedule P – Analysis of Losses and Loss Expenses in the National Association of Insurance Commissioners (NAIC).

Esta base de datos contiene información de los reportes hechos por las principales líneas personales y comerciales de todas las aseguradoras de daños y perjuicios que están registradas en los Estados Unidos. En la base de datos existen seis líneas de aseguramiento las cuales son:

- private passenger auto liability/medical
- commercial auto/truck liability/medical
- workers' compensation
- medical malpractice – claims made
- other liability – occurrence
- product liability – occurrence

Para este trabajo solo se dispondrá de la base de datos relacionada con los seguros de las malas prácticas médicas (medical malpractice – claims made) con el fin de explorar solo un caso de aseguramiento. Es importante esclarecer desde un principio los limitantes de este tipo de proyectos y las suposiciones en los cuales se cimienta. El problema de estimación de provisiones posee el limitante o desventaja de que solo es preciso cuando los patrones del desarrollo de pérdidas del pasado continúan funcionando en el futuro, por esta razón cuando existen cambios en las operaciones de aseguramiento como los cambios en los tiempos de liquidación de las reclamaciones o los cambios en las prácticas de reserva de provisiones puede suceder que los métodos a desarrollar no produzcan estimaciones precisas de aprovisionamiento si no se hacen los respectivos ajustes al modelo a desarrollar por lo cual se tiene que este tipo de problemas son muy sensibles a los cambios y pueden ser impropios para ciertas líneas de negocio muy volátiles.

1.3 OBJETIVOS EN LA CIENCIA DE DATOS

Una vez establecidos los objetivos del problema es necesario identificar como estos objetivos se ven traducidos en las labores de la ciencia de datos, es decir determinar cual es el propósito del análisis de datos que se desarrollara durante el proyecto. De esta manera, en el tratamiento de los datos se fijan los siguientes objetivos:

- Identificar los patrones que permitan estimar la reserva de provisiones en determinados tipos de seguros a partir de la base de datos histórica CAS Loss Reserve expuesta anteriormente. (Modelo de forecasting)
- Establecer medidas de desempeño que permitan valorar el modelo a desarrollar en la base de datos del proyecto y que sean compatibles con los triángulos de pérdida que provee la base de datos.

1.4 PLAN DE PROYECTO

Para la consecución de los objetivos de este proyecto en primer lugar es necesario iniciar una fase de entendimiento de los datos para comprender la distribución y la información que recopila la base de datos. Posteriormente, dado que los datos ya se encuentran ordenados, se inicia la fase de modelado en donde se busca poder determinar las reservas de provisiones a partir de registros históricos consignados en los triángulos de pérdida usando el método de Chain-ladder. Una vez establecido el modelo se procede a fijar una medida de desempeño adecuada que permita evaluar la capacidad predictiva de nuestro modelo con los datos de evaluación extraídos de la base de datos. Una vez completas todas las fases y después de un proceso de depuración y corrección de los métodos empleados y los resultados obtenidos se procede a la fase de despliegue del modelo.